

ASSESSING THE EFFECTIVENESS OF ENVIRONMENTAL TEST PROGRAMS

Steve Cornford, Mark Gibbel and Tim Larson
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive, Pasadena, CA 91109

BIOGRAPHY

Dr. Steve Cornford graduated from UC Berkeley with undergraduate degrees in Mathematics and Physics and received his doctorate in Physics from Texas A&M University in 1992. He has focused his efforts at JPL to establish a quantitative basis for performing reliability assurance including environmental testing. Among other activities, he is currently the Payload Reliability Assurance Program Manager for a number of NASA-funded research efforts at JPL and the Group Supervisor of the Reliability Technology Group.

Mr. Mark Gibbel graduated with a Bachelor of Science in Mechanical Engineering from California State Polytechnic University at Pomona and has been involved in design and test of advanced electronic packaging since 1977, when he developed the thermal models for Hewlett Packard's flip-chip technology. He has dedicated his professional career to failure physics based test and verification methodology development and implementation, with a particular interest in space flight electronics. Since 1985, he has been a member of the Reliability Technology at the Jet Propulsion Laboratory where he has innovated and been the lead for several research projects within NASA's Test Effectiveness Program.

Mr. Tim Larson received a Bachelor's Degree in Mechanical Engineering in 1985 from CSULA. He is the Thermal and Dynamics Environmental Compatibility and Reliability Lead in the Reliability Engineering Office at JPL. He manages portions of the NASA Code QE sponsored Test Effectiveness Program at JPL in addition to supporting several flight projects.

KEYWORDS

Test Effectiveness, Environmental Testing, Test Improvement, Optimization, Test Tailoring, Metrics Evaluations, Corporate Culture, Anomaly Data

ABSTRACT

NASA's Code QE Test Effectiveness Program is funding a series of applied research activities focused on utilizing the principles of physics and engineering of failure along with those of engineering economics to assess and improve the value added by various validation and verification activities, with the primary emphasis being on

testing. Presented herein is a review of the methodology and data sources required for successful metrics evaluations and the utility to, and feasibility of implementing at, other organizations. Also covered are brief examples of test metrics which enable useful evaluations as well as a brief summary of the relevant failure modes, effectiveness of the test control parameters on screening effectiveness, and applicability of the metrics defined.

INTRODUCTION

The NASA Administrator has challenged all the NASA Centers to do things "Faster, Better and Cheaper".¹ He has also been quoted as saying "If you can't measure it, you can't manage it!". At the Jet Propulsion Laboratory, we have found that the second of these two statements illuminates the means to achieve the first. JPL is already well down the path to Faster, Better and Cheaper spacecraft and this has been accomplished by making extensive use of tailoring, especially test programs. The fabrication of spacecraft which are state-of-the-art, highly reliable and one-of-a-kind presents significant challenges. One of the principal challenges has been to leverage the knowledge gained during the test program for one spacecraft to that of another. An even bigger challenge facing the industry has been in finding ways of collecting and analyzing data from a variety of sources in a way that enables the portability of findings across commercial and aerospace industries, and from high volume to ultra-low volume industries.

We have met these challenges by combining the Physics of Failure approach² with the engineering judgment inherent in ultra-low volume builds. Various aspects of this methodology have been described previously^{3,4,5,6}. This general methodology is referred to as *Physics and Engineering of Failure*. The general approach can be summarized by considering the various flaws (design, materials, or process) to be caught, or "screened", by the various detection and prevention activities which occur in the project life cycle. These include reviews, design rule implementation, process controls, analyses, and testing. By further dividing the product into its constituents one obtains finer resolution into the anomaly type (or "cause code") which allows one to perform higher resolution effectiveness evaluations. One continues this process until sufficient confidence in

X

the evaluation results is reached. This may include models (traditional physics of failure), developmental testing, or other detailed evaluations to determine "root cause". The main emphasis of this approach in the testing arena is the determination of the test stress which precipitated the failure mode. However, as discussed below, there is utility in evaluations at nearly every level of detail with certain caveats.

DATA SOURCES

source is
The most readily available source of data for internal metrics evaluations are industrial Failure Reporting And Corrective Action (FRACA) systems. High-level metrics studies⁴ often correlate the number of failures seen in a particular environmental test for a given project within a company. This results in valuable conclusions regarding test effectiveness for a specific technology and a variety of tests, with heritage and design/test/QC processes specific to that company. This can generate very useful data for trend studies to determine the general effectiveness of a specific test within the context of a single organization, with a fairly constant technology mix, design heritage, and corporate culture.⁷ However, this level of information is not always easily extrapolated to new technologies, or transferable across industries or even to different companies within an industry, due to the differences in the way PACTs* are applied in the development, design, build, and test process. This has resulted in either ignoring everyone else's data or issuing guidelines which overlook the differences in corporate cultures.^{8,9}

However, as stated above, implementation of a "root cause" methodology has allowed us to use industrial data when evaluations occur at the root cause level. External collaborative efforts have taught us that useful evaluations can be performed almost regardless of the particular corporate FRACA system. Usage of this external data has resulted in a number of applicable findings including the utility of thermal cycling¹⁰, relative cost data, and the role of thermal ramp rates and cold-biased functional testing⁶, to name a few. It is an interesting confirmation of the root cause approach that the high effectiveness of cold-biased functional testing is apparent in both JPL's ultra-low volume spacecraft product and a high-volume computer product.

Although the corporate FRACA systems provide valuable information, there is also a wealth of data available from internal process capability studies, Tiger Team problem solving, and design of experiments. These studies are typically quite focused and expensive, but if you have it, use it! This also points to keeping mind that useful

documentation of the interim findings may eliminate the need for future detailed evaluations. Another advantage of the methodology being described is that it enables identification of those areas where additional data generated by such studies would have the most pay-off, and what data is needed from ongoing data gathering activities.

Another important finding from utilizing commercial FRACA databases is that the comments entered by technicians, even though generally terse, in most cases, are insightful enough to allow physics of failure evaluations to be made. We did find, however, that in many cases, this data was going unused and represented an untapped wealth of knowledge. The message we are trying to convey is that it is not difficult to make significant improvements in your processes, but you need to utilize data relevant to your processes and that this data probably already exists.

In some cases, the existing data collection process could be modified to enable the collection of additional data that enables more accurate assessments. In particular, this data could include the specific test stress and operational conditions, time of failure, total operating time accumulated on the hardware, the history of tests that a particular hardware element has experienced, and the specific test stresses at the time of failure. This data collection process modification should be focused on the specific areas where more data is needed to perform those evaluations that are of the most importance to a particular organization.

Armed with useful data, one can now perform correlations which tie the test stress to a physical failure mechanism. Along with the knowledge of what activities (PACTs) were performed previously on the hardware and the operational history of the hardware, the correlation can now tie failure mechanisms to various test conditions. This enables the transfer of knowledge between technologies, between industries, and between companies with different corporate cultures.

At JPL, we are trying to leverage external, as well as internal, data to the greatest extent possible and several working groups have been established including the NASA Test Effectiveness Working Group and industrial collaborations to achieve this mutually beneficial process of sharing data and findings. If you are interested in participating, please contact one of the authors at the addresses provided in this paper.

* PACTs= Preventative measures, Analyses, process Controls and Tests.

? in ?

UTILITY AND APPLICABILITY OF METRICS STUDIES

The advantages of this root cause approach are that: 1) the data obtained from one spacecraft development program can be applied to another, 2) industry data can be leveraged ("physical failure mechanisms don't know which industry they occur in!") and 3) higher confidence in the overall effectiveness of a detection and prevention program can be obtained. This allows the conservative approaches of the past to be winnowed down to the subset which really gives the most "bang for the buck".

With the ever-growing capability and availability of advanced technologies, even previously large volume industries must begin looking at root cause, and thinking like lower-volume industries do, to remain competitive in tomorrow's marketplace.

This paper is intended to convey the principles behind these evaluations by examining their application to a specific JPL project and motivate other companies to utilize their existing data in a similar fashion. Ultimately, the pooling of this knowledge across organizational boundaries will benefit all parties involved.

TOOLS

The tools required to successfully perform these evaluations are nothing special. We have utilized the relational database in which our Problem/Failure Reports (P/FRs) are stored and commercially available spreadsheets to do all of the evaluations in this paper. While JPL's P/FR system happens to be in FoxPro®, others are in usage (such as Sybase®, Microsoft Access® and DOORS® to name a few). Most of the available databases come with the requisite search engines.

Commercially available spreadsheets (such as Microsoft Excel®) are also useful tools for setting up relational matrices (such as failure mode types versus where in the process they were detected) and performing subsequent correlations and trend analyses.

One of the most important tools is open communication with the product design, test and support personnel. The insight provided by the people directly involved with the hardware is not always captured in an anomaly log process.

Another tool used is often referred to as a 'waterfall chart'. This is a plot of the detection and prevention activities versus the failure modes being detected or prevented. This plot is depicted graphically for visualization purposes in Figure 1, but in practice is usually captured in a spreadsheet format. Note that in this figure, the solid lines represent failure modes which are potentially present, while the dashed lines represent

"escapes" from a particular activity. This chart is used to convey three ideas. Firstly, there may be failure modes for which nothing is being done (the far right line) and we better do something about them! Secondly, be aware of the danger of eliminating those steps which represent the last, or only, chance to detect a failure mode (the line in the middle which is detected only by "System Performance"). Lastly, focus efforts on reducing those steps which are redundant with previous steps (arrows which line up above many boxes). Which case is which can be determined by performing metrics evaluations which characterize the flaw distributions present at various stages of development. This is the part which is most corporate culture sensitive but is also the most readily available from existing FRACA systems. *had*

A subtle aspect of implementing the various processes to detect or prevent failure modes is the fact that these processes may themselves introduce new failure modes. For example, a rework process may result in solder splashes or part overstressing. One may use the waterfall chart to illustrate this introduction of new failure modes by adding new lines emerging from a given activity box.

JET PROPULSION LABORATORY PROCESS

The process for effectiveness evaluations at NASA's Jet Propulsion Laboratory is depicted in Figure 2. There are always a number of projects in various phases of development, and the questions asked regarding test improvements depend on this phase. For example, in the very early proposal phase, project needs include identification of an overall test program with associated schedule and resource requirements. As the design progresses, questions become more specific and begin to focus on test tailoring, including test substitution, and test phasing. As hardware is fabricated, the questions become very specific and concern specific test levels and risk/consequences associated with deviations or waivers on specific tests. The test effectiveness team gathers up these Needs and Questions and then tries to identify areas where the questions are recurring and can be made generic.

UNIVERSAL METRICS

Some of the most useful metrics have been found to be generic and do not require constraint to specific tests or other activities. For example, the associated costs, the induced damage, the schedule impacts and effectiveness by general failure mode type are important metrics for any test. These universal metrics can result in correlation of defect distributions and quantities to:

- Product timelines
- Hardware types
- Failure Mode types

- Test (or PACT) types

The utility and fidelity of evaluations at these levels can be greatly enhanced by having detailed information regarding the failure modes or mechanisms. While these metric types are universal, the specific results may only be applicable to a particular corporate culture unless the root causes are evaluated.

As an example, if one does a high level analysis which identifies workmanship as being principally caught during a thermal test, what does one do with the answer? First of all, one can determine the relative effectiveness of the thermal test on finding all of these workmanship defects by examining prior and subsequent process steps and evaluating the associated failure mode distributions. This identifies escapes or the introduction of new workmanship defects. Secondly, one can generically act on the knowledge that workmanship problems account for a particular percentage of the hardware defects. However, this then only identifies a problem and does not yet help with the solution.

If however, it is known which type(s) of workmanship defect is being detected (more information regarding the failure modes) one can do two additional things: 1) Identify and improve the specific portion of the fabrication process which produces these failure modes, and 2) Correlate these failure mechanisms with the specific part of the thermal test which was finding them (i.e. was it the associated functional testing, the hot "burn-in", the transition to cold temperatures, etc.).

MEASUREMENT PARAMETERS

To correlate the effectiveness of a test to a particular failure mode one must examine the test itself to find the stresses it is inducing. For each test, there are a collection of test parameters, or test considerations, which are decided upon and implemented when the test is run. For example, in a thermal test one must decide the temperature levels, ramp rates, etc. but one must also decide how to configure, instrument and functionally interrogate the test. We refer to this collection of test implementation decisions as *test parameters*. One can now consider the influence each of these test parameters has on the effectiveness of the individual test versus particular failure modes.

Going back to our workmanship example, one would use information regarding where in the thermal test profile the failure occurred to determine the test parameters that were responsible. If the failure mode occurred while doing functional testing during a cold-to-hot transition, one can now begin to consider the interplay between changing temperatures and Coefficient of Thermal Expansion (CTE) expansion, temperature distributions or uniqueness

of the functional test. This correlation with actual failure modes is what allows one to identify that 1) a new functional test was being done and therefore its perceptiveness may have had nothing to do with the thermal test, or 2) the transition resulted in relative expansion between two materials exposing a defect or 3) the transition resulted in temperature gradients which caused two functional elements which were performing at the "ragged edge" of their performance margin to produce the anomaly. e
XX

The process continues until the user has sufficient information to focus improvement efforts or make "on-the-fly" tailoring decisions.

These types of detailed correlations are made possible by the existence of a methodology, a useful and accessible anomaly recording system, and the knowledge present in the heads of the various discipline engineers.

SUMMARY

This paper has described the need for performing metrics evaluations to achieve the goal of doing more with less. The need, and current availability, of the necessary data has been discussed. The process for utilizing the available data has been described which is a "top-down" approach that ends when one has sufficient information to identify required process improvements or tailoring options.

With this new methodology in place, the internal information gathering systems and the ability to leverage this data across institutions, we may all be able to do Faster, Better and Cheaper even better.

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology under a contract with the National Aeronautics and Space Administration through Code Q.

TO CONTACT THE AUTHORS

mark.gibbel@jpl.nasa.gov
steven.cornford@jpl.nasa.gov
tim.larson@jpl.nasa.gov

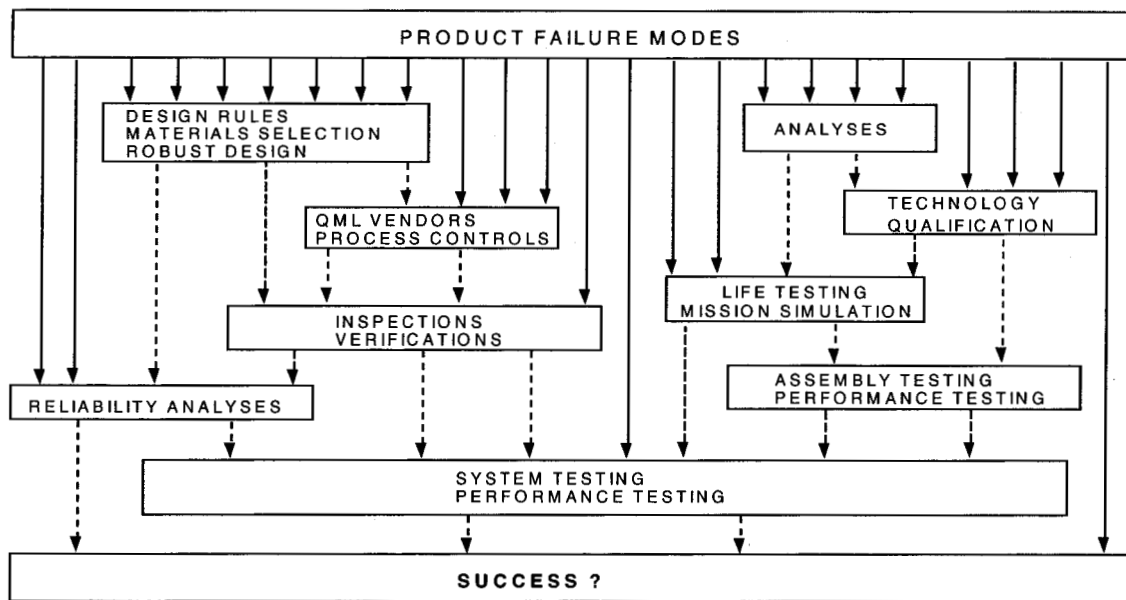


Figure 1 The 'waterfall chart' graphically depicts the role of various detection and prevention activities in screening out the failure modes.

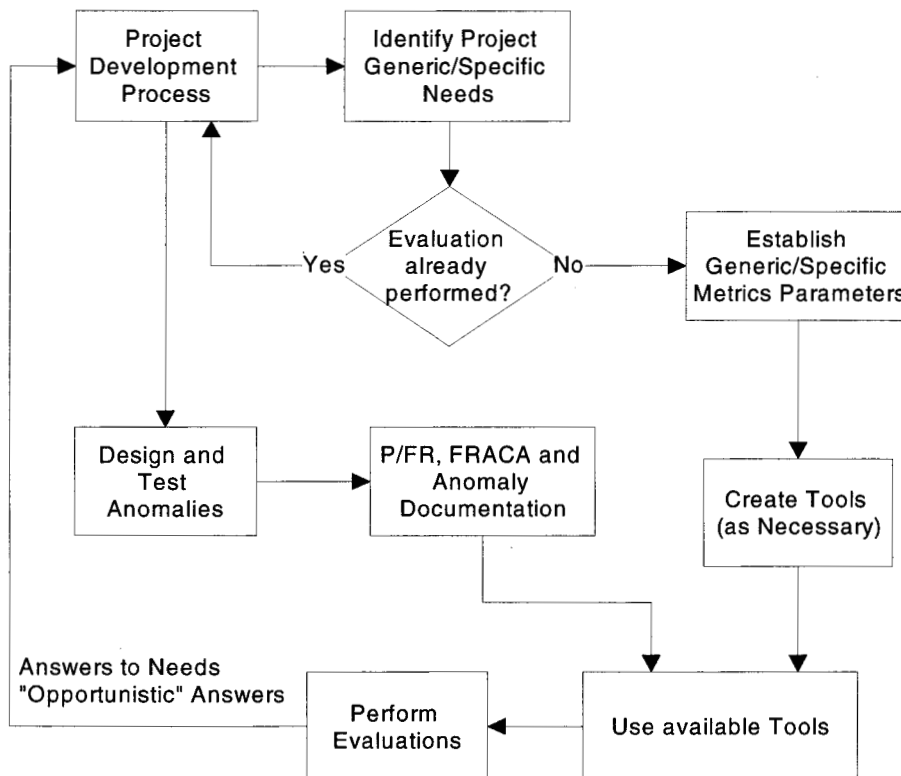


Figure 2 Flow chart for process of developing metrics questions, utilizing available tools and resources and providing answers.

REFERENCES

- ¹ 1997 NASA Strategic Plan
- ² P.Lall, M.Pecht, An Integrated Physics-of-Failure Approach to Reliability Assessment, Advances in Electronic Packaging ASME EEP-Vol.4-1, 1993.
- ³ M. Gibbel, 'Fundamentals of Failure Physics – Overview', IEEE Components, Hybrids and Manufacturing Technology, 1992
- ⁴ S. Cornford & M. Gibbel 'Methodology for Physics & Engineering of Reliable Products', WESCON 96 Proceedings, Anaheim, CA, October 22-24, 1996
- ⁵ S. Cornford, 'Defect Detection and Prevention', Proceedings of the IES Annual Technical Meeting, 1995
- ⁶ Environmental Stress Screening 2000 Final Report, National Center Manufacturing Sciences Report 0006RE97, April 1997
- ⁷ 'Environmental Test Effectiveness Analysis Reports', JPL D-11295, Rev C; JPL Internal document
- ⁸ For example, "Environmental Stress Screening – A Tutorial", Institute of Environmental Sciences, Mount Prospect, IL, 1985
- ⁹ For example, MIL-STD-1540D "DOD Standard Practice, Product Verification Requirements for Launch, Upper Stage and Space Vehicles", 1997 and MIL-HDBK 340A (Vols. I and II) "Test Requirements for Launch, Upper Stage and Space Vehicles – Baselines and Application Guidelines", 1997, US Air Force Documents
- ¹⁰ S. Cornford and P. Plumb, 'Thermal Cycle and Thermal Dwell Report', 15th Aerospace Testing Seminar, El Segundo, CA, October 1994